Immigrant Children's School Performance and Immigration Costs: Evidence from Spain*

Facundo Albornoz

Antonio Cabrales

Paula Calvo

Esther Hauk

March 2018

Abstract

This note provides evidence on how immigration costs affect school performance of immigrant children exploiting the information provided by the CDI; a standardized exam for all students enrolled in the last year of Primary education in the Madrid region. For a given socio-economic background and parent characteristics, school performance of immigrant children improves with parental immigration costs.

JEL-Classification: I20, I21

Key-words: school performance, immigration, parental involvement, immigration costs

*Hauk acknowledges financial support from the Spanish Ministry of Economy and Competitiveness through the second Severo Ochoa Program for Centers of Excellence in R&D and through CICYT project number ECO2015-67171-P (MINECO/FEDER) and from the government of Catalonia through AUGUR 2017 SGR 1571. Email addresses: facundo.albornoz@nottingham.ac.uk (Albornoz), a.cabrales@ucl.ac.uk (Cabrales), paula.calvo@yale.edu (Calvo) and esther.hauk@iae.csic.es (Hauk).

1 Introduction

In this note, we provide evidence on the relationship between immigration costs and immigrant children's educational achievement. In general, few studies have explored the determinants of school achievement of immigrant children. Those studies focus on the impact of children's characteristics, such as age at migration (Corak, 2011), language proficiency (Dustmann, Machin and Schönberg, 2010; Dustmann and Glitz, 2011) or whether one or both parents are immigrants (Ohinata and van Ours, 2012).

However, the effect of immigration costs as a determinant of school performance has been overlooked. Clearly, immigration costs affect immigration decisions and selection. Moreover, when people migrate as a family the future opportunities of immigrant children are likely to weigh in the migration decision. As argued by Albornoz, Cabrales and Hauk (2017), the overall benefits from migration will then depend on parental motivation and aspirations. One aspect of immigrant selection is therefore parental involvement, which establishes a link between immigration cost and the school performance of immigrant children. When parental selection improves in immigration costs, the school performance of immigrant children should increase in parental immigration cost. To the best of our knowledge, this is the first paper establishing this empirical association.

Our results are based on a standardized evaluation of the universe of students (native and migrants) at the region of Madrid (Spain). The case of Madrid is interesting because of the high proportion of immigrant population (17% in 2010), the variety of origins (with the 10 top immigration origins explaining 70% of total immigration) and the fact that immigration is a relatively new phenomenon (immigrants comprised only a 0.5% of total population in 1981).[1] We find that immigrant children perform better if their parents faced higher emigration costs. This result is robust to controls for family and country of origin characteristics. Therefore, the positive association between immigration costs and immigrant children's school performance cannot be explained by selection of migrants in skills or education levels or the quality of education in the origin country.

[1]Information provided by the Observatorio Permanente de la Inmigración and Spanish Statistical Office.

2 Empirical Hypothesis and Data

Albornoz, Cabrales and Hauk (2017) propose a theoretical model linking immigration and the school system. In this model, student outcomes are determined endogenously as a result of the interplay between different families (immigrants and natives) and the school system. The effect of immigration on schooling depends on parents' characteristics, such as wages and skills, but also on parental motivation, which refers to parents' concerns about their child's educational achievement and their perceived value / importance of their involvement for their child's success. This model identifies conditions under which, for any skill level, immigrants are positively selected in parental motivation. When highly motivated parents enjoy relatively higher benefits from emigrating, selection in parental motivation improves with higher emigration costs. This implies our empirical hypothesis: *for a given host country, immigrant children who perform better are those whose parents faced higher immigration costs.*

We test this hypothesis exploiting the information provided by the *Prueba de Conocimientos y Destrezas Indispensables* (CDI), a standardized exam taken at the last year of primary education in all primary schools in the Madrid region. The exam consists of two sections of 45 minutes each, Language and Maths, and it is conducted in Spanish for all students, independently of the school type. The CDI has the advantage of providing precise information about immigrant students, which stems from a questionnaire on student's individual and family characteristics, filled out by each student before taking the exam. This allows us to identify the country of origin of the student and their parents, and the student's arrival age to Spain.

We focus on students who took the CDI in 2010. We associate the immigrant group of a student by the father's country of origin (focusing on mothers does not affect the results), independently of where the student was born. Our measure of performance is given by the student's total score in the CDI, standardized to the yearly mean.

Table 1 reports the average total score and the average Maths and Language scores of the CDI exam in 2010 for the ten largest migratory groups in Madrid. While immigrants perform on average worse than Spanish students, the performance of immigrant students varies considerably across origins. Students whose fathers are migrants from Ecuador, Bolivia, Dominican Republic or Morocco are outperformed, on average, by those students whose fathers are Romanian, Polish or Russian.

We use the sample of immigrant students from the CDI to test whether variations in performance across different immigration groups are grounded in differences in immigration costs across immigrant's source countries. For

Table 1: Mean scores in CDI test for principal migratory groups.

Country of birth of the father	Total Score	Language	Mathematics	Sample Size
Ecuador	**20.00**	**12.34**	**7.65**	2,654
	(8.99)	(4.85)	(5.18)	
Morocco	**18.66**	**11.09**	**7.51**	1,215
	(10.15)	(5.81)	(5.37)	
Romania	**22.60**	**13.34**	**9.23**	1,195
	(9.35)	(4.90)	(5.47)	
Colombia	**21.60**	**13.00**	**8.68**	912
	(8.99)	(4.82)	(5.27)	
Peru	**23.10**	**13.71**	**9.41**	866
	(8.90)	(4.47)	(5.43)	
Dominican Republic	**16.47**	**10.36**	**6.12**	563
	(9.58)	(5.69)	(4.98)	
Bolivia	**20.53**	**12.59**	**7.94**	520
	(9.15)	(4.94)	(5.30)	
China	**20.72**	**9.57**	**11.08**	387
	(10.71)	(6.71)	(5.27)	
Poland	**25.63**	**14.98**	**10.69**	201
	(8.51)	(3.99)	(5.47)	
Russia	**27.54**	**15.57**	**12.43**	22
	(9.66)	(4.17)	(6.02)	
Average for immigrants	**21.88**	**12.97**	**8.89**	11,765
	(9.63)	(5.16)	(5.54)	
Average for all students	**25.88**	**14.91**	**10.95**	52,246
	(8.94)	(4.38)	(5.46)	

Note: Standard deviations in parenthesis.

children i, in migratory group k and school s, we estimate the following model by OLS:

$$Score_{iks} = ImmigrationCosts_k'\alpha + x_i'\beta + z_k'\delta + \gamma_s + \mu_{iks}. \tag{1}$$

Immigration Costs is a vector of variables that proxy for the difficulty of migrating from origin country k to Madrid. This vector contains an indicator variable of whether country k has Spanish as the official language, *Common Language*$_k$; the log of bilateral distance between k and Madrid, $ln(Distance)_k$[2]; and the share of the migratory group k over total immigration, *Proportion of Migrants*$_k$, computed from the CDI. The implicit assumption is that immigration costs increase with distance and decrease with the share of the migratory group and when countries share a common language. We also consider an alternative measure of migratory fixed costs developed by Grogger and Hanson (2011), $MigratoryCostsHG_k$, which aims to capture both direct monetary costs of migration, and the monetary value of psychological costs and source specific immigration policies imposed by Spain.

The vector x_i' includes variables to control for socioeconomic background of students, which are potential determinants of student performance, and are constructed from the CDI. Following Anghel and Cabrales (2014) these controls include the child's arrival age to Spain (*Arrival Age*$_i$), variables capturing the mother's education level, the highest parental occupation, the family composition, the age and gender of the student, and the age when child started formal education. The vector z_i' controls for characteristics of the country of origin which are not directly associated with the cost of migration to Spain, but which might affect the performance of students.

Our main results control by the logarithm of the GDP per capita of the country of origin in 2010, *ln(GDPpc)*$_k$. We use the logarithm of the GDP per capita since this variable is available for most countries in our sample and is correlated with other measures of performance coming from standardize international assessments, such as PISA (correlation=0.77), PIRLS (correlation=0.42) and TIMSS (correlation=0.46).[3] In the Appendix, we control instead for the quality of the school system at origin using country performance in these international standardized assessments. Since international exams are not implemented in every country, our sample is significantly reduced in these specifications. Although this reduces the power of our estimation, our results

[2]Bilateral distances are calculated as the distance between the biggest city in the country of origin and Madrid, weighted for the share of the city in the overall country population. This variable is obtained from CEPII website: http://www.cepii.fr/.

[3]GDP per capita was obtained from the World Development Indicators database.

Table 2: Dependent Variable: Standardized Aggregate Score in the CDI exam.

	(1)	(2)	(3)	
$CommonLanguage_k$	-0.158***	-0.320***		
	(0.037)	(0.041)		
$ln(Distance)_k$	0.066***	0.132***		
	(0.020)	(0.019)		
$Immigrant\ Share_k$	-0.806***	-0.294**		
	(0.136)	(0.138)		
$MigratoryCostsHG_k$			0.004***	
			(0.001)	
Socioeconomic controls (x_i)	No	Yes	Yes	
Characteristics of origin (z_i)	No	Yes	Yes	
School FE (γ_s)	Yes	Yes	Yes	
Observations	11,755	10,037	8,468	
R-squared	0.273	0.414	0.420	
Robust standard errors clustered by school and country or origin of the father in parenthesis.				*** p<0.01, ** p<0.05, * p<0.1

are robust to these controls.

Finally, γ_s captures school fixed effects. Since scores may be associated with specific characteristics of the migratory group and the schools in which they may concentrate, we cluster errors to allow within source country and within school correlation.

3 Results

Table 2 reports our main results. In column (1), we estimate performance against our migration costs variables, without additional controls. The coefficient associated with $CommonLanguage_k$ is both negative and significant. Coming from a Spanish speaking country has a negative effect on student performance. The coefficient associated to $ln(Distance)_k$ is positive and significant. Arguably, a larger distance from Madrid is likely to involve higher migration costs, which could give place to a selection of more motivated parents. Finally, the coefficient associated to the variable $Proportion\ of\ Migrants_k$, is negative and significant, supporting the hypothesis that a more important network of migrants reduces schooling outcomes, as migration costs decrease.

In column (2) we include additional determinants of student performance.[4] A main concern is that the variables considered as proxies of immigration costs

[4]The coefficients for these additional controls are not reported, but are available upon request.

may be correlated with other characteristics of the country of origin which might affect performance. We control for the logarithm of GDP in 2010 of the source country. While this variable is positive and significant, the effect of our immigration costs variables remain qualitatively unchanged. We also add variables to control for socioeconomic and cultural status of students and their families. These variables have the expected signs, consistently with Anghel and Cabrales (2014). Reassuringly, adding these controls does not affect the signs and significance of our variables of interest.

Results in column (2) are quantitatively relevant. Sharing the official language reduces the score of the students in 0.32 standard deviations, a marginal effect of 11%. Taking two countries as Ecuador and Poland this variable explains, *ceteris paribus*, half of the difference in the average CDI score between these countries. An increase of 1% in distance, increases the score in the CDI exam in 0.0013 standard deviations. In the case of two countries as Poland and Morocco (with distances to Madrid of 2,293 and 762 kilometers, respectively), this is equivalent 1.3 points, explaining 19% of the difference in the average score. For the importance of the migrant network, an increase in 1 percentage point in the share of migrants negatively affects the aggregate score in 0.003 standard deviations. In the case of Ecuador and Romania (with a proportion of migrants of 22.7% and 10.3%, respectively), this explains 13% of the difference in average scores existing between both countries.

Finally, in column (3) we use the direct measure of immigration costs proposed by Grogger and Hanson (2011), and include the same controls as in column (2). Again, the results show a clear positive relationship between migration costs and student performance.[5] Table A1 in the appendix shows that our findings are qualitatively similar when we use the scores in Mathematics and Language separately, when we exclude countries with a similar level of development than Spain,[6] and when we control for the quality of the school system at origin.

4 Conclusion

We document a positive association between immigration costs and immigrant children's school performance. This result is relevant to understand achievement gaps between immigration groups, which is a key issue of public policy debates about integration of immigrants and persistent income inequality

[5]Estimations of migratory fixed costs are not available for some relevant source countries as Morocco and Argentina, which have been excluded in this specification.

[6]These are the OECD members as of December 31st, 1989.

across different ethnic or nationality groups. Our main finding is also useful to design targeted support for learning of immigrant children. Importantly, the fact we uncover in this note highlights once more how essential is to understand the relevance of parental involvement in the learning process.

References

[1] Albornoz, F., Cabrales A. and E. Hauk (2017) "Immigration and the School System" Economic Theory.

[2] Anghel, B., and A. Cabrales (2014) "The determinants of success in primary education in Spain," Journal of public programs and policy evaluation 2, 22-53

[3] Corak, M. (2011). "Age at immigration and the education outcomes of children," Working Paper 6072, Institute for the Study of Labor (IZA)

[4] Dustmann, C., and A. Glitz (2011). "Migration and education," Handbook of the Economics of Education, 4, 327-439.

[5] Dustmann, C., S. Machin and U. Schönberg (2010). "Ethnicity and educational achievement in compulsory schooling," The Economic Journal, 120, 272-297

[6] Grogger, J., and G. Hanson (2011): "Income maximization and the selection and sorting of international migrants," Journal of Development Economics, 95, 42-57

[7] Ohinata, A., and J. C. van Ours (2012). "Young immigrant children and their educational attainment," Economics Letters, 116(3), 288-290

A Robustness checks

Table A1 report the results of estimating similar specifications to those displayed in column (2) of table 2. Columns (1) and (2) in table A1 disaggregates total scores by Mathematics and Language. Results are qualitatively similar to what we obtain for the aggregate score. Columns (3) and (4) replicate results in columns (2) and (3) of table 2, but excluding those countries with similar level of development than Spain. Clearly, our results hold after this re-sampling. Finally, results in columns (5) and (6) control for performance in PIRLS and PISA assessment respectively. In both cases, the signs of our variables of interest remain unchanged, but we lose power in the estimations.

Table A1: Dependent Variable: Standardized Scores in the CDI exam.

	(1) Math	(2) Language	(3) Total	(4) Total	(5) Total	(6) Total
$CommonLanguage_k$	-0.373***	-0.192***	-0.377***		-0.259**	-0.147*
	(0.040)	(0.047)	(0.045)		(0.124)	(0.076)
$ln(Distance)_k$	0.172***	0.059**	0.152***		0.132**	0.012
	(0.019)	(0.023)	(0.022)		(0.060)	(0.039)
$Immigrant\ Share_k$	-0.369***	-0.194	-0.238*		-0.070	-0.458
	(0.131)	(0.152)	(0.141)		(0.672)	(0.621)
$MigratoryCostsHG_k$				0.004***		
				(0.001)		
Socioeconomic controls (x_i)	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics of origin (z_i)	ln(GDP)	ln(GDP)	ln(GDP)	ln(GDP)	PIRLS	PISA
School FE (γ_s)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,143	10,242	9,120	7,552	4,062	4,612
R-squared	0.378	0.371	0.403	0.404	0.510	0.477

Results in columns (3) and (4) exclude countries with similar development level than Spain.
Robust standard errors clustered by school and country of origin of the father in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1